

**Projet de rapport  
du groupe de travail du CNIS  
Clarté et accessibilité des données du  
système statistique public**

**VERSION du 22 novembre 2001**



## **Préambule**

### **Rappel du mandat du groupe de travail**

Paris, le 3 octobre 2000 n°532/D130

#### **Mandat du groupe de travail consacré à la clarté et à l'accessibilité des informations issues du système statistique public**

---

Le CNIS est fondé, dans l'esprit de la loi du 7 juin 1951, à s'assurer de la qualité des informations statistiques rendues publiques. La qualité statistique est couramment comprise dans un sens strict comme la précision, mesurée par l'analyse des différents types d'erreurs, liées ou non à l'échantillonnage.

Cependant il est reconnu au niveau international et en particulier européen que les composantes de la qualité des statistiques vont bien au delà de la précision.

Ainsi les critères de qualité comportent-ils la pertinence (dont le CNIS s'assure constamment dans son examen des programmes et des opérations statistiques) et également l'accessibilité et la clarté ainsi que l'actualité et la ponctualité.

Le Bureau du CNIS a été saisi de façon récurrente, et notamment en 1996 et en 1999, de cette question par plusieurs organisations syndicales.

Le débat public comporte de nombreux enjeux : les données chiffrées disponibles pour les différentes parties prenantes peuvent conditionner aussi bien le déroulement de négociations sociales que des transferts financiers vitaux pour certains secteurs de l'économie. Il est indispensable, dans ces conditions, que les acteurs économiques et sociaux sachent précisément quelle est la qualité des informations statistiques qu'ils utilisent.

Le Bureau du CNIS a décidé, lors de sa réunion du 16 mars 2000, de la création du groupe de travail "consacré à la lisibilité des données chiffrées et aux procédures conduisant à leur publication".

Ce groupe s'attachera à évaluer la situation actuelle au regard de :

1. l'accessibilité et à la clarté des statistiques issues du système statistique public ;
2. l'actualité et à la ponctualité de ces statistiques.

Il élaborera des recommandations en vue d'atteindre :

- la meilleure adéquation possible des niveaux d'explication et d'analyse des statistiques en fonction des publics y compris en matière d'évaluation des politiques publiques, lorsque celle-ci repose sur des statistiques issues du système statistique public ;
- le développement de pratiques au sein du système statistique qui garantissent la ponctualité et l'actualité de la mise à disposition des données par exemple par le respect d'un calendrier de mise à disposition annoncé à l'avance ;

- une meilleure mobilisation des technologies de communication, en vue d'un accès facilité aux dossiers statistiques fondamentaux.

Ces recommandations devront s'apprécier au regard des démarches statistiques mises en œuvre dans différents pays de l'Europe et de l'OCDE.

Le groupe de travail rendra ses conclusions en septembre 2001.

Il sera présidé par M. Jean-Pierre BOMPARD, membre du Bureau du CNIS. Son rapporteur sera M. Michel JACOD, directeur de la diffusion et de l'action régionale à l'INSEE.

Le groupe sera ouvert aux membres du Conseil (qui pourront désigner un représentant) et pourra associer des personnalités qualifiées.

J.M. DELARUE

### **Rappel des limites de la réflexion**

Le Groupe de travail n'avait pas mandat à s'interroger sur l'opportunité des opérations statistiques, ses réflexions ont donc porté sur l'accès à des données collectées et/ou traitées d'après les programmes statistiques des membres du système statistique public.<sup>1</sup>

Les analyses et propositions ci-dessous peuvent toutefois être étendues au-delà du champ actuellement couvert par le CNIS dans ses deux composantes : enquêtes soumises au label et opérations de statistique administrative. En particulier, ont été citées l'abondante production d'information chiffrée des collectivités locales et celle des organismes paritaires de sécurité sociale qui pourraient s'inspirer des recommandations ci-dessous.

Par ailleurs, la plupart des administrations et organismes publics produisent de l'information chiffrée pour un objectif de gestion interne ; l'absence de traitement de type statistique (précision des concepts mesurés, protocole défini d'observation, comparabilité dans le temps et dans l'espace, redressements d'anomalies, ...) peut conduire à ne pas les intégrer dans cette réflexion, même s'il s'agit à l'évidence de « données publiques » au sens du projet de loi sur la société de l'information. Cette question du champ des opérations concernées est restée délicate tout au long du travail du Groupe. Persuadé que l'accès à l'information doit être le plus ouvert possible, le Groupe estime cependant que les recommandations qu'il suggère peuvent et devraient, dans leur esprit du moins, trouver à s'appliquer sur un champ plus large que le seul programme statistique du CNIS. Il lui semble de plus que l'intégration de ces préoccupations dans la labellisation « statistique publique » pourrait apparaître comme un avantage utile à certains producteurs de données et ainsi les encourager à le mériter.

Le Groupe de travail n'avait pas mandat à s'intéresser directement aux moyens et méthodes permettant l'évaluation des politiques publiques et plus généralement l'action publique ou privée sauf en ce qui relève de la mise à la disposition des acteurs par le système statistique des informations qui leur sont nécessaires pour leurs fins.

En particulier, la plupart des services producteurs cultivent une compétence d'étude voire de recherche dont les résultats ne relèvent pas du champ de ce Groupe de travail consacré aux conditions de mise à disposition des données chiffrées, ni d'ailleurs à proprement parler du CNIS. Ces activités se situent dans un domaine marqué par le pluralisme des approches et, de ce fait, la mise à disposition de ces analyses n'obéit pas aux mêmes nécessités de régulation que celle qu'impose la situation de monopole dans laquelle est organisée la statistique publique.

Enfin, faute d'une connaissance suffisante des dispositifs étrangers, le groupe de travail n'a pas eu la capacité de faire des comparaisons systématiques avec la situation à l'étranger ou dans des organisations internationales.

---

<sup>1</sup> Les débats ont bien sûr amené le groupe à aborder certaines lacunes dans la collecte des données, ces remarques seront transmises aux formations compétentes du CNIS.

## Nature des propositions

Le groupe de travail n'avait pas les ressources nécessaires pour réaliser une véritable « évaluation » de la situation actuelle en termes de clarté, d'accessibilité, d'actualité et de ponctualité des statistiques produites et mises à disposition par le système statistique public. Une telle investigation aurait nécessité un travail très important du fait de la diversité des thèmes étudiés par la statistique publique, des organisations qui en sont en charge, de la variété des opérations statistiques périodiques ou non, et des usages auxquels elles peuvent se prêter.

Par ailleurs les situations problématiques qui sont à l'origine de ce Groupe de travail (cf. mandat du Groupe) montrent à l'évidence que la situation peut être variée et changeante dans une organisation donnée. Il valait donc mieux qu'il se consacre à réfléchir aux risques de dysfonctionnements les plus notoires plutôt que de se livrer à une analyse systématique d'un existant dont la pérennité n'est justement pas acquise.

Le Groupe de travail s'est donc appuyé sur les cas particuliers qui avaient amené le bureau du CNIS à décider de sa création ainsi que sur les remarques des participants à ses travaux qui ont largement élargi les perspectives de départ.

Sans mésestimer la valeur des principes déontologiques qui sous-tendent d'ailleurs largement les propositions auxquelles il a abouti, le Groupe de travail a pris le parti de faire des recommandations en général d'ordre procédural.

Le rapport comprend deux parties, un constat de questions qui se posent dans le champ du mandat et quelques suggestions en regard, suivi d'une liste récapitulative de ces recommandations.

## 1. Des risques de dysfonctionnement

### 1.1. Des collectes ou traitements statistiques qui ne donnent pas ou plus lieu à diffusion

Le cas n'est pas fréquent. Il peut bien sûr s'expliquer en tout ou partie par des difficultés de méthode ou des questions budgétaires ou pratiques de tous ordres ; les opérations statistiques ne sont en effet pas plus que d'autres opérations techniques à l'abri de ce genre de risques. L'expérience montre aussi que d'autres considérations plus politiques peuvent intervenir ou être soupçonnées d'intervenir. Dans tous les cas, le risque d'amalgame avec de la rétention d'information est grand si une information n'est pas faite auprès des commanditaires sûrement mais aussi des autres utilisateurs potentiels.

L'annonce a priori d'échéances de publication lors de l'inscription des opérations au programme statistique, l'examen annuel de la réalisation du programme statistique par le CNIS constituent des opportunités qu'il conviendrait d'utiliser plus systématiquement. Toutefois ce rythme annuel n'est pas toujours satisfaisant.

Lorsque les enquêtes ou opérations sont à périodicité infra-annuelle, on peut penser que les utilisateurs sont alors suffisamment nombreux et attentifs aux sorties pour que l'absence d'une diffusion attendue soit d'une façon ou d'une autre manifeste et donc qu'une information soit donnée par le service producteur préalablement ou peu après l'échéance attendue ; cette situation ne peut qu'être améliorée par la publication régulière des échéances à venir comme c'est le cas des indicateurs économiques les plus courants.

La question ne se pose pas de la même manière s'agissant d'opérations aperiodiques ou totalement originales mais dont l'actualité est importante, par exemple parce que devant permettre l'évaluation de politiques publiques précises. La crédibilité d'un service statistique peut être entachée si une information n'est pas faite régulièrement sur les difficultés rencontrées quelle qu'en soit la nature assortie d'une mise à jour es échéances initialement prévues. L'annonce des échéances lors du lancement ou de la préparation d'une opération est actuellement une procédure fréquente si ce n'est générale, elle gagnerait donc à être doublée par une procédure de rappel de ces échéances ; cette procédure devrait être systématique et selon un rythme ajusté aux échéances initialement prévues pour chacune des opérations et en tout cas définis en dehors des enjeux de cette information. Le secrétariat du CNIS pourrait ainsi demander aux producteurs et afficher une telle mise à jour personnalisée des échéances annoncées. Cela viendrait compléter les informations fournies à la partie 6. *Publications* des avant-projets de programmes présentés au CNIS par les services

producteurs. On peut aussi penser que chaque producteur d'information statistique aurait à cœur d'offrir en parallèle cette information sur son site internet.

Concernant les opérations régulières mais à périodicité ou à échéance longues - plusieurs années par exemple entre deux enquêtes ou entre la collecte et les résultats - un enlèvement technique ou/et une éventuelle rétention d'information sont bien plus difficiles à mettre en évidence. Ce risque encore plus grand accroît la pertinence de la proposition ci-dessus.

## 1.2. Des collectes ou traitements statistiques dont la diffusion est tardive

Diffusion tardive peut s'entendre en deux sens :

- tardive par rapport aux enjeux de débat ou de développement (ex : résultats d'un recensement disponibles après la réflexion organisée autour de contrats de plan Etat-Région, résultats d'une enquête sur les modalités et effets du passage aux 35 H disponibles après la discussion parlementaire d'un projet de loi portant sur une 2<sup>e</sup> étape législative sur ce sujet, etc.) ;
- ou simplement tardive par rapport aux annonces faites initialement par le service producteur concerné.

Pour rester dans son mandat, le Groupe de travail n'a pas pris de positions sur le premier type de retards lorsqu'ils résultent de décisions claires fixant les dates d'opération statistique et donc les calendriers de diffusion correspondants. Il s'est plutôt penché sur les cas où une opération menée dans des délais adéquats ne se traduit pas par la mise à disposition de ses résultats en temps et heure.

En dehors bien sûr des opérations statistiques périodiques à forte visibilité – indicateurs économiques et principaux indicateurs sociaux, ce phénomène de retard est réputé bien plus fréquent que l'absence totale de diffusion évoquée précédemment.

Là encore, l'organisation d'une information en continu sur les échéances et leurs révisions paraît nécessaire.

Toutefois, au-delà de la réalité des faits, plusieurs phénomènes peuvent contribuer à accentuer, parfois outre-mesure, cette impression négative (les exemples ci-dessous sont choisis dans la diffusion de l'Insee mais chacun trouvera des exemples dans son champ de compétence de producteur ou d'utilisateur) :

- diffusion échelonnée des résultats du fait du volume considérable d'informations à organiser alors que seule les premières échéances ont été annoncées par le producteur ou mémorisées par les utilisateurs (recensement)
- toujours dans ce cas, retard effectif sur une partie du programme de diffusion qui efface les échéances tenues pour le reste du programme (*idem*)
- décalage perçu comme trop grand entre la sortie de 1<sup>ères</sup> analyses (de type 4 pages) et la disponibilité des données détaillées nécessaires à une évaluation pluraliste (enquête sociale)
- disponibilité de résultats dans un cercle restreint avant une diffusion générale (enquête handicap-incapacités-dépendance)

Le premier cas renvoie encore une fois à un plus grand effort de clarté dans l'information donnée sur les échéances. Quant au second cas, outre bien sûr les efforts toujours nécessaires sur le respect des échéances de production ou de réalisation des produits de diffusion en tout genre ou/et sur l'évaluation a priori de ces échéances, il met encore plus en évidence le besoin d'une communication suivie sur la situation. Là aussi, la gestion sur internet d'une telle information faciliterait la clarté de l'information initiale même complexe et sa mise à jour si nécessaire.

Les deux derniers cas méritent explication préalable.

La réalisation d'un minimum d'études est considérée par les statisticiens français comme le moyen ultime de validation des données de collecte après les processus de corrections d'anomalies individuelles ; il serait assurément dangereux de minimiser son rôle dans l'assurance de qualité donnée aux utilisateurs. Cette pratique courante depuis la création du système statistique français a d'ailleurs essaimé à l'étranger depuis les années 80. Au passage, ces « études », le plus souvent une simple lecture des résultats, constituent une manière de valoriser fortement le gisement d'information

car elles facilitent le portage à la connaissance des utilisateurs de la disponibilité des résultats de l'opération statistique en question (cf. 1.6).

Pour autant ce problème de décalage entre sortie des 1ères analyses et disponibilité des données détaillées ne peut toujours être simplement résolu et une recommandation de publication simultanée du 4 pages (ou équivalent) et des données détaillées serait sans doute irréaliste. En effet, les produits de diffusion qui permettent cette mise à disposition ne peuvent être constitués complètement avant la validation définitive des chiffres, même s'ils peuvent être en partie préparés ; par ailleurs, leur production (impression, création de bases de données, ...) répond souvent à d'autres impératifs d'organisation du travail que celle de la première publication rapide (4 pages) qui est directement l'affaire du statisticien responsable de l'opération.

Le dernier cas (accessibilité précoce impossible hors d'un cercle restreint) se rencontre surtout lors d'opérations statistiques innovantes (nouveau champ, nouvelle technique, nouvelle population enquêtée, ...). Pour de telles opérations, un groupe d'exploitation, d'analyse, ... quel qu'en soit le nom, est souvent constitué associant producteurs, chercheurs et parfois praticiens du domaine en question. Ces mêmes personnes et institutions ont souvent d'ailleurs largement contribué à la conception de l'opération. Pour ces opérations innovantes, la question de la validation des résultats se pose mais avec beaucoup plus d'acuité que dans le cas d'enquêtes ou opérations répétitives. L'exclusivité pendant un temps de l'accès aux données qui apparaissent ainsi réservées à un groupe particulier, se justifie pleinement par cette nécessité. Ceci peut bien sûr être expliqué mais impose de la clarté dans la constitution de tels groupes « privilégiés » : appel d'offres mais aussi sélection impartiale des participants sur leur engagement à participer à ce travail collectif, leurs qualités scientifiques dans le champ et la diversité des approches institutionnelles et scientifiques qu'ils peuvent représenter.

### 1.3. Des collectes ou traitements statistiques dont la diffusion n'est que partielle

Plusieurs raisons militent pour une mise à disposition la plus complète possible des données collectées ou traitées, sans s'arrêter comme cela est assez fréquent à la simple diffusion de synthèses chiffrées ou de résultats principaux :

- bien sûr, d'abord le développement des utilisations de l'information pour améliorer la connaissance et au passage rentabiliser au mieux l'investissement collectif puisque le coût de la collecte a été pris en charge par la collectivité ;
- plus particulièrement, pour satisfaire l'exigence démocratique de pluralisme des études et en particulier de l'évaluation des politiques publiques ;
- enfin pour permettre la reproductibilité des études et recherches qu'elles soient menées en interne par les services producteurs ou menées en externe.

Sous la contrainte des règles de confidentialité juridiquement définies (loi de 1951 modifiée, loi de 1978 sur l'informatique et les libertés, secrets divers prévus par les dispositions législatives ou réglementaires applicables notamment à certains fichiers administratifs), le Groupe de travail recommande que l'ensemble ou le sous-ensemble juridiquement acceptable des données détail ou micro-données qui ont pu être collectées ou traitées dans des procédures statistiques soit rendu accessible par le service producteur dans des conditions organisées<sup>2</sup>.

Compte tenu des formes que prennent actuellement les règles de secret applicables, le Groupe de travail est bien conscient que si cette recommandation a une portée assez générale dans le domaine des données sur les personnes, en particulier des enquêtes sur échantillon, cette dernière est plus restreinte dans le domaine des données sur les entreprises.

Sur certains domaines toutefois, les services statistiques produisent régulièrement des indicateurs synthétisant de tels ensembles de données selon une méthodologie affichée et généralement bien connue des spécialistes<sup>3</sup> sans que ces ensembles de données eux-mêmes puissent facilement être rendus disponibles (ex comptabilité nationale pour des raisons de complexité, indices des prix pour prévenir toute initiative susceptible de fausser l'instrument de mesure plutôt que d'agir sur la réalité à

<sup>2</sup> Seules des conditions organisées de mise à disposition permettent un accès équitable des utilisateurs aux données publiques (cf. 1.7).

<sup>3</sup> Même si elle est évidemment discutable et d'ailleurs remise en chantier de temps en temps, de manière institutionnelle ou à l'initiative ou avec la participation active du système statistique (ex : modification des catégories de chômeurs, changement de base d'indices ou de la comptabilité nationale).

mesurer, etc.). Dans ces cas vraisemblablement peu nombreux, le groupe de travail est d'avis de maintenir le statu quo.

Dans l'esprit d'amener à la plus grande ouverture possible aux fichiers de données détaillées, le Groupe suggère de considérer que cette ouverture doit être la règle pour les opérations statistiques sauf dispense de diffusion des données détaillées qui pourrait être étudiée et accordée par une instance ad hoc : pour les enquêtes statistiques lors du passage en comité du label du CNIS ; pour les autres opérations statistiques, une procédure devrait être imaginée dans cet objectif.

#### 1.4. Des collectes ou traitements statistiques insuffisamment documentée pour leur diffusion

Il est rare qu'une source statistique ne soit pas documentée par le service producteur au moins pour son usage interne, et peu fréquent qu'elle ne le soit pas lorsqu'elle est mise à disposition des utilisateurs externes. Toutefois cette situation jugée globalement assez bonne pourrait encore être améliorée.

Dès lors que la diffusion sur internet des sources statistiques<sup>4</sup> se développe, la nécessité de cette documentation technique (en gros : la méthodologie de la source, la signification des variables, la fiabilité et les limites d'utilisation des données) apparaît criante, au point qu'en étant un peu optimiste, il pourrait suffire de recommander la mise sur internet, cette dernière équivalant à organiser une pression de la demande suffisante pour que les services producteurs atteignent à terme l'objectif prescrit.

Au-delà de cette documentation technique par source, le groupe de travail estime qu'un effort doit être fait pour permettre aux utilisateurs d'établir des liens avec les sources disponibles sur le même sujet. C'est la logique des « *pour en savoir plus* » que l'on retrouve sous une forme ou une autre dans les publications statistiques habituelles. Là encore, l'outil internet peut favoriser cette mise en relation par l'utilisateur et enrichir sa démarche en indiquant plus facilement les liens à faire et en proposant à l'utilisateur de les mettre en oeuvre directement. Dans la situation actuelle, certains des sites internet des services producteurs renvoient de façon générale aux sites des autres producteurs intervenant sur leur champ ; assez souvent sont aussi présentés des renvois bibliographiques (sous forme de liens internet ou non) vers d'autres publications.

Il est aussi possible de trouver de telles informations sur le site CNIS, notamment à la rubrique « Les enquêtes ». On est pourtant très loin d'un système convivial voire simplement efficace pour l'utilisateur peu au fait de l'organisation administrative un peu complexe de la statistique publique française.

Une mention particulière doit être faite concernant les sources européennes ou plus généralement internationales. Sans mésestimer les problèmes de comparabilité, le groupe estime qu'un renvoi correctement documenté à des sources d'origine internationale lors de la mise à disposition des données françaises correspondantes devrait devenir une préoccupation des producteurs français.

Dès lors que les données statistiques sont de plus en plus mobilisées dans le domaine de la recherche, il serait également utile de faire bénéficier les utilisateurs des exploitations secondaires qu'en ont fait des chercheurs. Ces dernières peuvent apporter des éclairages complémentaires à ceux que donnent les statisticiens de leur propre production, suggérer des utilisations plus variées des données et renforcer de façon très intéressante la documentation native des sources y compris parce que certaines faiblesses des données sont ainsi parfois mises en évidence. La mise en place du Centre Quételet<sup>5</sup> suggère au groupe de travail qu'une organisation devrait être mise en place pour démultiplier en faveur des utilisateurs externes au monde de la recherche les efforts de circulation d'information sur les exploitations secondaires des sources que ne manquera pas de faire ce nouveau Centre à la suite du LASMAS et estime que, pour être efficace, cette organisation devra s'appuyer sur les informateurs primaires que sont les services producteurs.

---

<sup>4</sup> Diffusion de synthèses statistiques sous forme d'indicateurs et de 4 pages d'analyse, mais aussi de données détaillées sous forme de tableaux agrégés intégrés ou non dans des bases de données généralistes et sous forme de micro-données non confidentielles ; techniquement ces derniers points sont encore délicats et coûteux à organiser mais ne devraient pas le rester très longtemps, et en attendant un accès à une maquette peut sans trop de difficulté être organisé.

<sup>5</sup> explication sur le champ d'action du Centre et renvoi au décret fondateur .....

### 1.5. Des collectes ou traitements statistiques diffusés selon des programmes de diffusion inadéquats

Les supports et modes techniques de mise à disposition (papier, disquettes, cédéroms, internet, serveurs FTP, ... mais aussi l'organisation en collections, la présentation des données sous forme de bases de données directement réutilisables ou simplement en lecture, etc.) jouent évidemment un rôle très important dans l'accessibilité, l'actualité et la ponctualité des données, voire dans leur clarté. C'est vraisemblablement sur toute la gamme des supports et moyens techniques disponibles que doivent jouer les statisticiens pour rencontrer au mieux les besoins et les capacités d'exploitation de leurs utilisateurs.

Sans devoir devenir le média unique à court terme, internet est particulièrement adapté à la nature des informations produites par le système statistique public. Son développement rend la situation fortement évolutive et la complexité de sa mise en oeuvre<sup>6</sup> explique pour partie la diversité des offres actuelles des différents services producteurs.

Dans ces conditions évolutives et assez techniques, le groupe de travail s'est contenté de réflexions générales sur les procédures à mettre en œuvre pour renforcer les chances d'atteindre les objectifs indiqués dans son mandat sans construire des recommandations techniques précises d'application forcément trop contingente.

Deux thèmes sont apparus intéressants sur ce plan :

#### *a) organiser le dispositif de diffusion pour ouvrir l'accès aux données*

Comme toute activité, l'effort de diffusion reste limité par des considérations pratiques. Ces dernières font l'objet de décisions prises essentiellement par les services producteurs en fonction de leur compréhension de la demande effective et sous la contrainte des moyens disponibles. Quelles qu'elles soient, les modalités techniques de mise à disposition ainsi décidées ne peuvent ni anticiper toutes les demandes potentielles pour des raisons de connaissance a priori de ces demandes ni même envisager de les satisfaire sous forme de produits préconçus du moins si on souhaite rester dans une enveloppe raisonnable de coûts.

Si le principe évoqué plus haut d'une accessibilité généralisée des données<sup>7</sup> est établi, il convient alors d'organiser leur diffusion en mettant en place une ou des procédures permettant l'accès dans des conditions raisonnables de coût à cette diversité de données.

L'ouverture de l'accès aux fichiers détail dans le respect des contraintes légales et réglementaires de confidentialité est une manière de faire qui renvoie de façon simple aux utilisateurs le soin de mettre en œuvre les demandes originales qui leur conviennent. Lorsque ceci n'est pas possible, une procédure d'accès « à distance », permettant le traitement sur mesure attendu, devrait être mise en place de manière généralisée.

#### *b) développer la proximité avec les utilisateurs*

Une concertation active avec les utilisateurs potentiels sur les modalités les plus adaptées de mise à disposition est souvent organisée autour des plus grandes opérations statistiques par le service producteur ou/et le CNIS. Concernant les opérations périodiques, une telle concertation pourrait avoir lieu de temps en temps, à l'occasion de l'un de leurs examens annuels par la formation concernée du CNIS. Les opérations nouvelles quant à elles ne donnent pas forcément lieu à la construction d'un dispositif de diffusion particulier ; le plus souvent et c'est évidemment souhaitable à la fois pour des raisons d'économie pour les producteurs et de lisibilité pour les utilisateurs, leur diffusion s'insère dans un dispositif général ; l'examen périodique dans le cadre du CNIS des dispositifs généraux de diffusion des services producteurs pourrait être l'occasion d'améliorer leur adéquation aux besoins et capacités des utilisateurs.

Toutefois de telles concertations ne produiront les fruits attendus que si elles s'appuient sur un réel travail préalable. Ce travail devrait comporter des « études de marché<sup>8</sup> » portant à la fois sur les

<sup>6</sup> Dans la situation actuelle, une véritable complexité apparaît dès lors qu'on dépasse la simple mise à disposition d'objets figés (ex : 4 pages, mais aussi textes plus consistants ou bases de données pré construites) ou de séries permanentes d'indicateurs. En particulier, donner à l'internaute la possibilité de se constituer lui-même sur mesure l'ensemble de données qui l'intéresse à partir d'un gisement d'informations de grande taille nécessite un investissement coûteux pour le producteur.

<sup>7</sup> Hors données confidentielles et cas de dispense explicite.

besoins et sur les capacités des utilisateurs dans leur diversité et des analyses de lectorat sur l'existant. Il devrait aussi comporter du côté des producteurs l'élaboration et la présentation de schémas de diffusion<sup>9</sup> globaux.

### 1.6. Des collectes ou traitements statistiques peu connus des utilisateurs potentiels

L'accessibilité des données suppose que soit connue leur existence et donc que les services producteurs constituent une offre qu'ils rendent ensuite visible en la portant auprès des utilisateurs potentiels.

#### *a) de bonnes pratiques bien répandues*

Les systèmes en place sont divers selon les sujets et selon les services producteurs, mais on retrouve de façon assez générale l'existence de publications légères de type 4 pages dont la fonction primaire – transmettre un message à l'intention d'un public a priori très vaste – atteint souvent dans les faits un objectif secondaire - faire connaître la disponibilité des résultats d'une opération statistique.

La bonne gestion d'une relation suivie avec les médias, généralistes ou spécialisés, est alors de nature à transformer cet exercice en en redoublant la portée.

#### *b) à développer par la participation à des réseaux*

Plus généralement, une telle relation suivie avec des intermédiaires proches des utilisateurs ou des réseaux d'utilisateurs va dans le même sens.

Dernier-né dans ce monde de réseaux, le Centre Quételet déjà cité plus haut apparaît très prometteur sur ce terrain puisqu'il devrait organiser la promotion de l'offre en données détaillées des services producteurs de statistiques auprès du public des chercheurs, public particulièrement intéressant pour la valorisation des opérations statistiques. Le groupe de travail recommande donc qu'une coopération étroite s'établisse progressivement avec cette nouvelle institution.

De la même manière, l'organisation de groupes d'exploitation autour d'enquêtes ou d'opérations nouvelles, si elle doit être conduite en prenant quelques précautions (cf. ci-dessus § 1.2), permet une publicité naturelle de ces opérations qu'il faut simplement veiller à ne pas limiter au cercle des participants directs.

Aller plus loin pour toucher directement les utilisateurs supposait, jusqu'à il y a peu, des actions de marketing assez coûteuses que peu de services producteurs faisaient l'effort de mener en dégageant les capacités nécessaires.

#### *c) et en mettant à profit les outils internet*

Internet ouvre là aussi des perspectives un peu nouvelles.

La plupart des services mettent d'ores et déjà sur internet les publications légères évoquées plus haut, de manière coordonnée avec la diffusion plus traditionnelle sur support papier ou fax. Mais cette bonne pratique reste relativement passive, son effet reste limité par la notoriété du site support, le groupe de travail suggère donc que d'autres possibilités d'internet soient mises à profit pour développer l'information sur l'existant statistique et développer ainsi l'accès aux données.

Internet devrait d'abord offrir la possibilité à un internaute de découvrir l'existant en accédant aux sites statistiques. Sur ce plan, l'organisation de la statistique décentralisée en France entre de nombreux acteurs se retrouve au niveau des sites internet réduisant ainsi fortement la visibilité de l'ensemble. De plus, la situation actuelle des sites de nombreux services producteurs de l'administration au sein du site de leur ministère n'est pas satisfaisante dans cet objectif, en effet l'organisation de ces derniers et leurs modes de navigation interne ne rendent pas ces rubriques statistiques accessibles de façon immédiate ou facile.

Internet devrait permettre aussi la recherche d'informations ciblées. Pour cela, encore faudrait-il que les sites statistiques soient référencés dans les principaux moteurs de recherche. La mise en œuvre d'une politique de référencement a un coût que les services producteurs ne peuvent sûrement pas

<sup>8</sup> Le terme d'études de marché est ici utilisé sans renvoyer à l'idée de tarification des données. Une étude de marché est le moyen de connaître les aspirations des clients sans que cela n'implique nécessairement une tarification du produit ou du service.

<sup>9</sup> De tels schémas préparés par les services producteurs devraient présenter une segmentation de leur offre inspirée par la connaissance qu'ils ont de leurs publics et l'analyse du champ à couvrir par les différents supports et « collections ».

tous assumer et le point précédent constitue une difficulté supplémentaire. Par ailleurs, le langage utilisé dans ces divers sites n'est en l'état actuel pas du tout normalisé.

Enfin internet peut être un moyen de communication avec les utilisateurs de la statistique dans un double sens :

- faire connaître les nouveautés par le moyen d'envoi d'informations personnalisées ou semi-personnalisées (« newsletter ») aux utilisateurs connus, mais cela suppose une approche « client » que peu d'organisations statistiques peuvent sans doute aujourd'hui envisager très sérieusement ;
- recevoir et traiter des demandes d'informations de la part de ces utilisateurs, sachant qu'une bonne part de ces demandes est naturellement transversale à l'organisation de la statistique publique en France, ce qui nécessite donc de s'appuyer sur une organisation et des outils ad hoc qui ne peuvent se limiter au champ de chacun des services producteurs.

Toutes ces perspectives seront peu ou prou indissociables dans un proche avenir et ce serait sans doute passer à côté des opportunités nouvelles que nous offre ce support que de cantonner l'internet statistique français à n'être qu'un ensemble de vitrines, en l'occurrence même pas juxtaposées<sup>10</sup> et pour certaines assez profondément cachées, et dans lequel les utilisateurs devraient faire l'effort de venir chercher ce qui leur convient.

Le groupe de travail est donc d'avis qu'une synergie forte devrait être recherchée entre les sites des services producteurs.

### **1.7. Des collectes ou traitements statistiques dont la diffusion est insuffisamment équitable**

Prendre en compte les besoins de connaissance de notre société et de pluralisme des approches conduit à remettre en cause des dispositifs de diffusion qui n'organisent pas un accès équitable aux données. Sans forcément d'intentions a priori, c'est néanmoins fréquemment le cas lorsque l'activité de diffusion d'un service producteur n'est pas organisée en tant que telle, faute de moyens pour l'ensemble du service ou de priorité donnée à cette activité.

En effet, la seule réalisation de produits de diffusion (annuaires, ouvrages de référence, produits électroniques, site internet) ne suffit pas toujours à satisfaire les besoins des utilisateurs, notamment parce qu'ils n'ouvrent assez souvent l'accès qu'à un sous-ensemble limité des données normalement disponibles.

Lorsque l'existence des informations statistiques est bien portée à la connaissance des utilisateurs potentiels, l'absence d'une organisation pour la diffusion des données (produits de diffusion standard mais surtout données détaillées, réponse aux questions) conduit à ce que l'accès aux données se fait par contact personnel avec le responsable direct de l'opération. Lorsque ce dernier peut se rendre disponible, les avantages de ce contact direct pour l'utilisateur particulier et aussi d'ailleurs pour le responsable direct sont indéniables ; mais, cette démarche trouve vite ses limites dans le nombre d'utilisateurs ainsi traités et met en place des relations favorisées entre ceux qui connaissent et les autres. Outre son caractère inéquitable, cette façon de faire ne conduit évidemment pas à multiplier l'accès à l'information.

## **2. liste des recommandations**

A l'issue de cette analyse, le groupe de travail souhaite émettre une série de recommandations. Ces recommandations sont de plusieurs ordres :

- les six premières, un peu procédurales, ne visent qu'à confirmer un mouvement déjà largement engagé par les services producteurs eux-mêmes et aussi dans certains cas par certaines formations du CNIS. Elles consistent à renforcer l'information sur l'information, en particulier sur les échéances, les contenus et les formes attendues de mise à disposition ;

<sup>10</sup> plusieurs vitrines peuvent se trouver sises côte à côte dans une « galerie marchande » et voir ainsi leur potentiel d'information de leurs clients tout à fait démultiplié (encore une fois, ces mots marchands sont indépendants du caractère payant ou gratuit de l'accès à l'information).

- les trois suivantes posent le principe de la complétude de l'information qui doit être mise à disposition et indique des pistes pour rendre effectif ce principe ;
- les trois dernières élargissent le point de vue en suggérant la création d'une véritable fonction de diffusion avec tous les composants requises au niveau de chaque service producteur et aussi à celui de l'ensemble du système statistique public, en indiquant la nécessité et les moyens d'une plus grande ouverture à l'Europe et à la recherche.

1. *Les indicateurs publiés régulièrement devraient faire l'objet d'un calendrier affiché par le service producteur pour une période glissante adaptée à leur périodicité (ex : pour des indicateurs mensuels, échéances fermes à 1 mois, révisables pour le trimestre suivant), la durée de cette période pouvant être progressivement allongée pour les nouveaux indicateurs entrant dans cette procédure.*

responsable : les services producteurs dans le cadre du CNIS

2. *L'inscription des nouvelles opérations au programme statistique devrait être assortie de l'annonce a priori de l'échéance de sortie des résultats.*

responsable : les services producteurs dans le cadre du CNIS

3. *Une information devrait être mise à disposition en continu sur ces échéances et leurs révisions, quelles que soient les causes de ces dernières.*

responsable : les services producteurs dans le cadre du CNIS

4. *Le secrétariat du CNIS devrait être responsable de la gestion d'une procédure de vérification systématique de l'état des échéances selon un calendrier adapté au déroulement prévu de chaque opération. Cette procédure pourrait être : ...*

responsable : le CNIS

5. *Ces échéances devraient concerner le programme complet de mise à disposition et non pas seulement la 1<sup>ère</sup> publication prévue.*

responsable : les services producteurs dans le cadre du CNIS

6. *Les programmes de publication devraient le plus possible s'inscrire dans des schémas de diffusion cohérents définis par chaque producteur pour l'ensemble de ses activités. De tels schémas devraient être construits en s'appuyant sur une appréhension organisée des besoins et capacités des utilisateurs, et bien sûr des ressources des producteurs et être périodiquement débattus au sein du CNIS.*

responsable : les services producteurs et le CNIS

7. *L'ensemble (ou le sous-ensemble juridiquement diffusible) des données détail ou micro-données qui ont pu être collectées ou traitées au cours de procédures statistiques devrait être rendu accessible par les services producteurs dans des conditions organisées (information sur les fichiers disponibles, procédure d'accès ou d'acquisition).*

responsable : les services producteurs

8. *Le CNIS devrait seul pouvoir accorder pour une opération particulière une dispense de diffusion des données détaillées (fichier détail et/ou données finement agrégées et présentées sous forme de bases de données) : le comité du label dans son champ de compétence (enquêtes statistiques), une organisation à définir pour les autres opérations.*

responsable : le CNIS

9. *Pour les sources dont la mise à disposition sous forme de publications et de produits standard (dont si possible les données détaillées évoquées plus haut) serait insuffisante, les services producteurs devraient proposer une ou des procédures permettant l'accès sur mesure à ces données.*

responsable : les services producteurs

10. *Pour organiser une promotion effective, une documentation de bon niveau dans son fond mais aussi dans sa présentation, une mise à disposition des données efficace et équitable, une communication autour de ces données sur les échéances mais aussi avec les médias et d'autres intermédiaires, et enfin pour assurer des réponses de qualité aux demandes des utilisateurs, les services producteurs devraient chacun organiser une activité de diffusion en tant que telle, clairement distincte de l'activité de production et d'étude et venant la compléter.*

responsable : les services producteurs

11. *La documentation d'une source ou de données devrait le plus systématiquement comprendre des liens avec les sources disponibles sur le même sujet, notamment les sources européennes ou internationales comparables ; ces liens devraient autant que faire se peut être proposés de façon contextuelle aux utilisateurs internautes.*

responsable : les services producteurs

12. *Les services producteurs et le Centre Quételet devraient organiser leur coopération pour mieux porter à la connaissance des milieux de la recherche l'existence des opérations statistiques, pour faire bénéficier les opérations statistiques nouvelles de compétences de chercheurs dans des conditions claires et évitant toute rétention d'information dans un cercle limité, et enfin pour faire bénéficier les autres utilisateurs des données des exploitations secondaires menées par des chercheurs.*

responsable : les services producteurs et le Centre Quételet

13. *Pour assurer dans des conditions optimales pour les utilisateurs l'accès via internet aux données statistiques, et pour faciliter le développement d'une présence effective, reconnue et visible de l'ensemble des sites ou rubriques de sites alimentés par les services producteurs, une forte synergie devrait être recherchée entre eux dans le développement d'un portail de service. Ce portail devrait offrir des services d'accès à l'information et d'orientation en pleine transparence de l'organisation administrative et promouvrait l'information disponible sans que la personnalité de chacun des services producteurs concernés ne soit remise en cause.*

responsable : les services producteurs.